The *noemen/heten* alternation: ongoing change in Colloquial Belgian Dutch

Dirk Speelman *(QLVL, KU Leuven)*

dirk.speelman@kuleuven.be

The linguistic variable

We discuss the position in Colloquial Belgian Dutch (CBD) of the variation exemplified in (1) and (2):

(1) Hij **heet** Tom. (*He**'s called** Tom.*)

(2) Hij **noemt** Tom. (*He**'s called** Tom.*)

In CBD, the verb *noemen*, which in standard Dutch means TO CALL, can also have the meaning *to be called* (2) so that it becomes a competitor for the (standard Dutch) verb *heten* (1).

The use of *noemen* is an atypical feature of CBD:

- Most CBD features originate from Brabantic dialects (AB on map below), whereas the *noemen* variant originates from the region Oost-Vlaanderen (OV on map below).
- Most CBD features are morpho-phonological or lexical, whereas the emergence of the *noemen* variant is a syntactic phenomenon.

Research questions

General:

RQ1 Which of our internal and external predictors correlate with the *noemen/heten* alternation, and how?

Specific:

RQ2 How does the distribution of the new variant (*noemen*) relate to that of most other CBD features, especially geographically?

RQ3 Which social groups lead the change?

RQ4 Do we find internal predictors that indicate that *noemen* emerged from a lexically and functionally specific expression/pattern?

Figure 1 : The levels of the predictor `region`

Materials

Data are from the face-to-face and telephone conversations in the Spoken Dutch Corpus (n=394, with 184 cases of *heten* and 210 cases of *noemen*).

Variables in models

- `variant`: levels `heten`, `noemen`
- `age`: speaker age, in years / `sex`: speaker sex, levels `F` and `M` / `region`: speaker birth region, levels `WV`, `OV`, `AB`, `LI` [see map] / `occup.type`: which education level does the speaker's occupation require? (levels `high`, `neutr`, `no.high`)
- `speaker`: speaker id, with 188 levels / `conver.type`: conversation type, with 2 levels
- `subj.cat`: category of subject, levels `dat`, `die`, `het`, `hijzij`, `PERS`, `THING` / `s.type`: sentence type, levels `D`, `Q` (resp. declarative and question) / `tc.left`: type of first character to the left, levels `vow`, `cons` / `collo`: collocate present? level `H`, `-`, `N` (resp. attracted to *heten*, neither, *noemen*).

Mixed-effects regression analysis

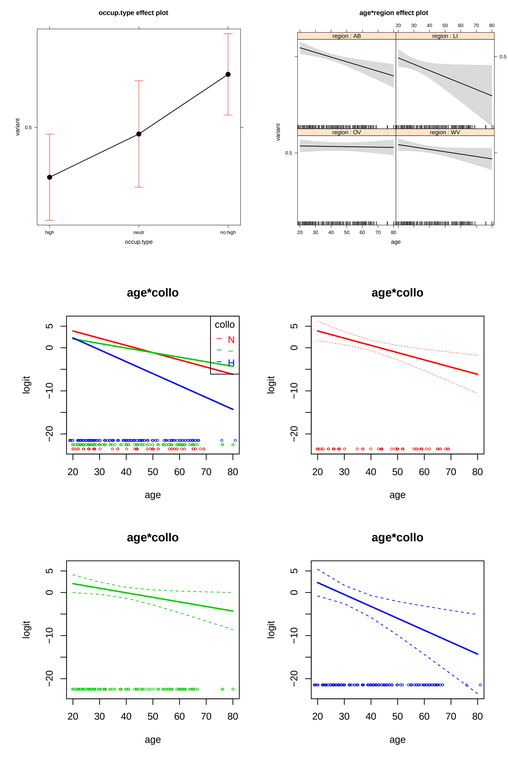

Figure 2 : The `glmer` model `variant ~ age*region + occup.type + age*collo + (1|speaker) + (1|conver.type)`, with, disregarding random effects, $C = 0.83$.

Summary report of regression model

- Strongest predictors are `age`, `region`, and `occup.type`
- In younger generation, use of *noemen* is spread all over Flanders, and now is strongest in `AB`
- Collocates of *noemen* 'make' younger generation use *noemen* [weak pattern]; Collocates of *heten* 'make' older generation use *heten*

Conditional inference tree

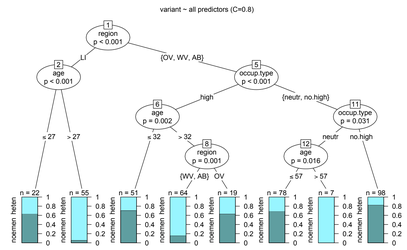

Figure 3 : Conditional inference tree (`ctree`) for `variant ~ all predictors`, with C=0.80

Random forest and variable importance

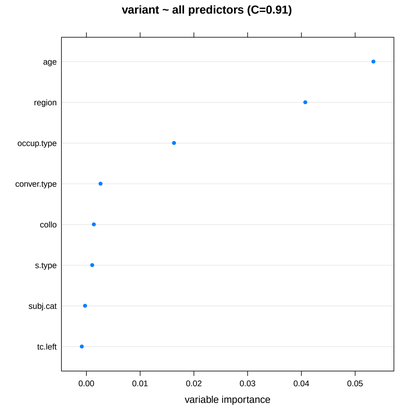

Figure 4 : Random forest (`cforest`) for `variant ~ all predictors`, with C=0.91

Summary report of inference tree

- Strongest predictors are `age`, `region`, and `occup.type`
- Overall, `occup.type=no.high` tends to correlate with high preference for *noemen*, but in `region=OV` there (also) is a high correlation of `occup.type=high` and preference for *noemen*

Summary report of random forest

- Strongest predictors are `age`, `region` and `occup.type`
- But there also are some (weaker) effects of internal predictors (`collo`, `s.type`)

Conclusions

General:

RQ1 External predictors have a much stronger effect than internal predictors, but there also are internal predictors at work, and they interact with the external predictors.

Specific:

RQ2 The (geographical) distribution of *noemen* has merged with that of the 'main body' of CBD

RQ3 But we have (weak) indications of different social groups leading the change in different regions

RQ4 There are (weak) lexically/functionally specific usages of *noemen*, but they are restricted to the younger users [There are stronger indications of lexically specific contexts where older speakers avoid *noemen*]

Next steps

- Other types of data will be needed to make a sharper distinction between apparent-time effects and age-grading
- Larger datasets will be needed to further study the weaker effects of the internal predictors (and their interactions with the external predictors)